



DETECTOR: structural information guided artifact detection for super-resolution fluorescence microscopy image

SHAN GAO,^{1,2,5}  FAN XU,^{3,5}  HONGJIA LI,^{1,2} FUDONG XUE,⁴
MINGSHU ZHANG,^{4,6} PINGYONG XU,^{2,4,7} AND FA ZHANG^{1,8}

¹High Performance Computer Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

²University of Chinese Academy of Sciences, Beijing 101408, China

³Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, USA

⁴Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

⁵These two authors contributed equally to this work

⁶mingshu1984@gmail.com

⁷pyxu@ibp.ac.cn

⁸zhangfa@ict.ac.cn

Abstract: Super-resolution fluorescence microscopy, with a spatial resolution beyond the diffraction limit of light, has become an indispensable tool to observe subcellular structures at a nanoscale level. To verify that the super-resolution images reflect the underlying structures of samples, the development of robust and reliable artifact detection methods has received widespread attention. However, the existing artifact detection methods are prone to report false alert artifacts because it relies on absolute intensity mismatch between the wide-field image and resolution rescaled super-resolution image. To solve this problem, we proposed DETECTOR, a structural information-guided artifact detection method for super-resolution images. It detects artifacts by computing the structural dissimilarity between the wide-field image and the resolution rescaled super-resolution image. To focus on structural similarity, we introduce a weight mask to weaken the influence of strong autofluorescence background and proposed a structural similarity index for super-resolution images, named MASK-SSIM. Simulations and experimental results demonstrated that compared with the state-of-the-art methods, DETECTOR has advantages in detecting structural artifacts in super-resolution images. It is especially suitable for wide-field images with strong autofluorescence background and super-resolution images of single molecule localization microscopy (SMLM). DETECTOR has extreme sensitivity to the weak signal region. Moreover, DETECTOR can guide data collection and parameter tuning during image reconstruction.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Super-resolution microscopy techniques, with the spatial resolution beyond the diffraction limit of light, provide unprecedented insights into subcellular structures at the nanoscale and have been widely used [1]. However, due to the complexity of instrument setup [2,3], imaging conditions [4,5] and reconstruction methods [6–8], super-resolution reconstructed images are prone to generate artifacts such as missing structure or structural sharpening. Although there is a lot of work to reduce the artifacts such as improving the reconstruction method [9], adjusting irradiation intensity and labeling density [4], it is difficult to get an artifact-free super-resolution image. These artifacts extremely hamper visual understanding of the structure of samples. To verify that the super-resolution images reflect the structure of the sample, the importance of developing robust and reliable artifact detection methods is widely concerned.

Generally, artifacts in super-resolution microscopy techniques refer to the differences between the underlying structures of samples and the reconstructed super-resolution image. However, due to the lack of ground truth structure, it is necessary to compare super-resolution images with alternative data. The conventional method is researcher-based detection, which relies on the prior knowledge of the expected structure or the benchmarking image reconstructed from other high-resolution imaging methods such as electron microscopy [10]. However, this manual strategy is subjective and demands immense human labor to cover all artifacts. Recently, Super-resolution QUantitative Image Rating and Reporting of Error Locations (SQUIRREL) [11], with reference to the wide-field image, transforms artifact detection problem into image similarity assessment problem. It detects artifacts by computing the absolute pixel difference of the wide-field image and the resolution-rescaled super-resolution image. However, this pixel-wise absolute mismatch computing is prone to report artifacts that do not exist (Supplement 1). There are three main reasons including autofluorescence background and noise of wide-field images, and the intensity inconsistency between the super-resolution image of single-molecule localization microscopy (SMLM) and the wide-field image [8]. Though SQUIRREL linearly transforms the intensity of super-resolution image to maximally match that of the wide-field image, this global intensity match is not suitable to a wide-field image that contains inhomogeneous background fluorescence (Fig. S1). Besides, the noise of wide-field image can affect the performance of SQUIRREL (Supplement 2). For super-resolution image of SMLM, there exist intensity information inconsistency between these two images. For a wide-field image, the intensity of each pixel depends on the label density of fluorescent molecules and the brightness of activated molecules. While, for super-resolution images reconstructed by SMLM, the intensity of each pixel is depend on the label density of fluorescent molecules and the blinking number of every single molecule. This intensity information inconsistency of the two images can lead to bias when calculating the error map depending on the intensity. (Fig. S2).

In this article, we propose a structural information guided artifact detection method (DETECTOR). It recognizes artifacts by computing the structural dissimilarity between the wide-field image and the degraded super-resolution image. To accurately compute structural information difference, our method has three key features. To ensures the resolution of degraded super-resolution images consistent with wide-field images, DETECTOR rescales the resolution of the super-resolution image with an actual point spread function (PSF) [12] which is measured from the optical imaging system. To make the imaging similarity assessment focus on the regions of structures, DETECTOR introduces a weight mask by extracting structural information from wide-field image. It is worth emphasizing that, trying background subtraction on wide-field image before processing with SQUIRREL is not at the same effect as structural information extraction (Supplement 3). Finally, based on the structural information, we proposed a structural similarity index for artifact detection in super-resolution images, named MASK-SSIM. The results show that DETECTOR can quantitatively detect artifacts in super-resolution images whose size is beyond the diffraction limit. DETECTOR can handle image that contains a strong autofluorescence background and super-resolution images of SMLM whose intensity information is inconsistent with that of wide-field image. DETECTOR has extreme sensitivity to the weak signal region and minor distorted reconstructed structures. Furthermore, DETECTOR can also help select reconstruction model and adjust model parameters.

2. Methods

2.1. DETECTOR framework

DETECTOR detects artifacts of super-resolution images by transforming the artifact detection problem into structural information guided image similarity assessment problem. The main idea is assuming that the reconstructed super-resolution image reflects the underlying structure of biological samples. After simulating the diffraction process on the super-resolution image, the

degraded super-resolution image and the corresponding wide-field image should be theoretically identical. However, considering the wide-field image contains strong autofluorescence background and the intensity information of super-resolution image of SMLM is not consistent with that of wide-field image, DETECTOR focuses on the structural similarity between these two images. Figure 1(A) shows the workflow of DETECTOR. The input of DETECTOR includes a super-resolution image, a wide-field image and an actual PSF. The output of DETECTOR contains an error map which shows the artifact distribution of super-resolution image and a similarity value of MASK-SSIM. There are three main modules in DETECTOR: (a) structural information extraction, (b) super-resolution image degradation, (c) image similarity assessment. In the structural information extraction module, DETECTOR extracts structural information from the wide-field image as a weight mask. In the super-resolution image degradation module, DETECTOR simulates the diffraction process and obtains the degraded super-resolution image. In the image similarity assessment module, DETECTOR proposes a new similarity index named MASK-SSIM. Based on the MASK-SSIM, DETECTOR identifies the artifacts by computing the similarity of the degraded super-resolution image and the wide-field image. The following is a detailed introduction of the three modules.

(a) Structural feature extraction. Here, we extract structural features of the wide-field image as the weight mask. The weight mask can help focus more on regions where present biological samples and filter out information that is regarded as less relevant during similarity assessment, such as fluorescence background. In this structural feature extraction module, we extract two key features to combine the weight mask. One is the edge feature that can preserve the important geometric properties of the biological sample; the other is the salient feature that distinguishes the target biological sample from the background. The details of how to extract edge feature and salient feature can see subsection 2.2.1 and subsection 2.2.2.

(b) Super-resolution image degradation. This module takes super-resolution image and an actual PSF which is measured from the optical imaging system as input and outputs resolution-rescaled super-resolution image. Before degradation, because wide-field image and super-resolution image may acquire from dual channels or different detectors, it is necessary to align these two images to avoid artifacts caused by position misalignment (see subsection 2.3.1). The accurate diffraction simulation is a key step to keeping the resolution consistency of the degraded super-resolution image and the wide-field image. Therefore, we adopted a Gaussian PSF model of the actual PSF to obtain a degraded super-resolution image (see subsection 2.3.2).

(c) Image similarity assessment. Here, we proposed a new image similarity assessment index based on structural similarity (SSIM) metric for super-resolution image, named MASK-SSIM (see subsection 2.4). This module computes the MASK-SSIM metric between the degraded super-resolution image and the wide-field image with a sliding window. It outputs an error map to present the artifact distribution of the super-resolution image and an overall similarity value score. In the error map, the intensity value of each pixel ranges from -1 to 1. It represents the similarity between the degraded super-resolution image and the wide-field image in the sliding window area. A higher pixel intensity value means a lower confidence of the reconstructed structure. Besides, the output similarity score is the mean intensity value of the error map.

2.2. Structural feature extraction

For existing super-resolution image artifact detection methods, directly computing the intensity difference between degraded super-resolution image and wide-field image is a simple and straightforward way. However, due to out of focus light and strong autofluorescence background signal, this kind of method is prone to report artifacts that do not exist. To solve the problem, we introduce the weight mask to make artifact detection focus on regions where exist biological structures and filters information that is regarded as less relevant. For wide-field images, many previous studies have proved that salient region and edge feature are significant in image

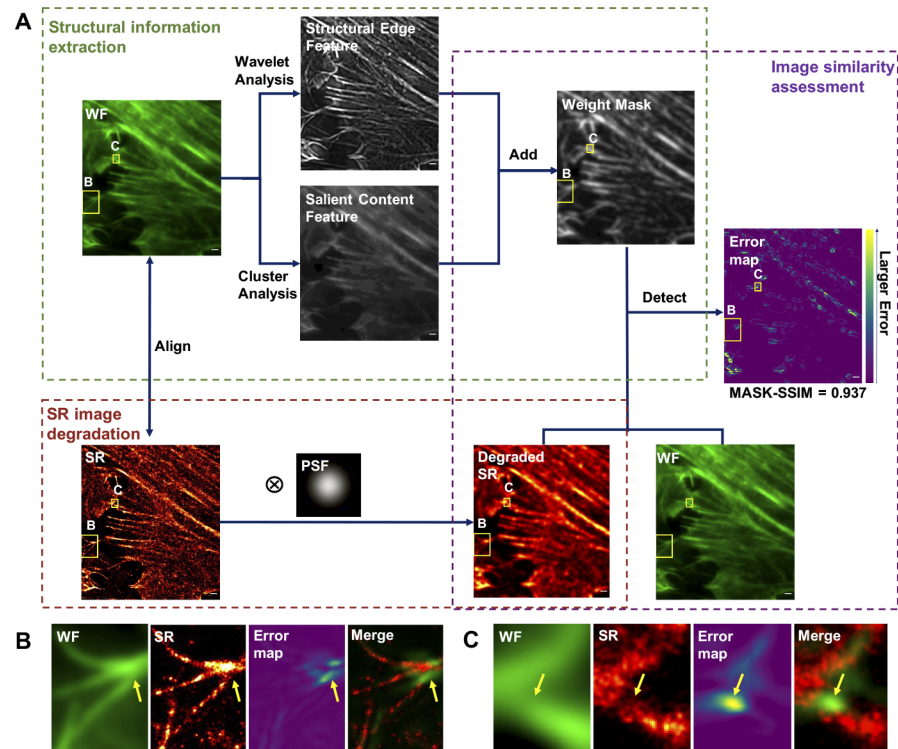


Fig. 1. Overview of the DETECTOR. (A) The workflow of DETECTOR. The wide-field (WF) image is colored in green and the super-resolution (SR) image, which is reconstructed by Photoactivated localization microscopy (PALM) [13] is colored in red. The subregions which are labeled with 'B' and 'C' is corresponding area for (B) and (C). The highlighted regions in the error map indicate inaccuracy reconstruction areas in the super-resolution image. A higher brightness corresponds to a more inconsistent between super-resolution image and wide-field image. The MASK-SSIM index is under the error map. The scale bar is 10um. (B) Details of the yellow box region B for figure A, from left to right: wide-field image, raw super-resolution image, corresponding error map, and a merged image with the super-resolution image in red and the error map in green. (C) Details of the yellow box region C for figure A, from left to right: wide-field image, raw super-resolution image, corresponding error map and a merged image with the super-resolution image in red and the error map in green. The line profiles of wide-field image and super-resolution image is shown in Fig. S5, which shown the mismatch distance between these two images.

understanding and analysis [14–17]. Based on these two features, we design a weight mask to emphasize the contribution of different image content to the image similarity assessment. For the salient region, we adopt mean-shift clustering to recognize regions where exist biological structures. For structural edge, we adopt wavelet analysis to extract edge features. To obtain a weight mask from these two features, considering the intensity scale of these images is different, we first normalized the brightness value of the feature images to 0-255. Then we obtained the mask image by adding these two images. To prevent the pixel intensity of the mask image out of bounds during sum, our mask image is 32-bit. In further similarity computation, we have normalized mask image to 0-1. The details of feature extracting by wavelet analysis and mean-shift clustering are Section 2.2.1 and Section 2.2.2.

2.2.1. Edge feature based on "à trous" wavelet transform

Edge feature extraction is considered an essential step in many computer vision tasks such as image segmentation, object recognition, and image classification [18]. As one of the multi-resolution decomposition methods, the "à trous" wavelet transform [19] can decompose the original image into an approximate image and a detailed image at a specific scale. Since it can analyze signals at multi scales, "à trous" wavelet transform can accurately extract edges of biological sample with different size objects. Accordingly, we applied "à trous" wavelet transform to extract edge features from the wide-field image.

Here, we assume the input wide-field image C_0 contains $N \times N$ pixels. In "à trous", C_0 can be decomposed into J ($J = \log_2 N + 1$) scale approximation images with the scale function f_l . For each scale $s = 2^j$ ($1 \leq j \leq J$), the spatial detail between the image C_s and C_{s-1} is minutia signal, which is generally called wavelet plane w_j . The edge feature of wide-field is the second wavelet plane. To obtain the w_2 , we choose the B3 cubic spline function [20] as the initial scale function. We first convolve the C_0 with the initial kernel k_0 (Eq. (3)) and get the image C_1 . Then insert zero between every two items in k_0 and get new kernel k_1 (Eq. (4)). The zero insertion in the kernel helps to subtract details of different scale information. C_2 is convolved with the new kernel giving the image C_1 . Then we subtract C_1 and C_2 to obtain the edge features.

$$C_{l+1}(i,j) = C_l(i,j) \otimes f_l(x,y) \quad (1)$$

$$w_{l+1} = C_l(i,j) - C_{l+1}(i,j), 0 < k \leq J \quad (2)$$

$$k_0 = \begin{pmatrix} \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{3}{128} & \frac{3}{32} & \frac{1}{64} & \frac{3}{32} & \frac{3}{128} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \end{pmatrix} \quad (3)$$

$$k_1 = \begin{pmatrix} \frac{1}{256} & 0 & \frac{1}{64} & 0 & \frac{3}{128} & 0 & \frac{1}{64} & 0 & \frac{1}{256} \\ \frac{1}{64} & 0 & \frac{1}{16} & 0 & \frac{3}{32} & 0 & \frac{1}{16} & 0 & \frac{1}{64} \\ \frac{3}{128} & 0 & \frac{3}{32} & 0 & \frac{1}{64} & 0 & \frac{3}{32} & 0 & \frac{3}{128} \\ \frac{1}{64} & 0 & \frac{1}{16} & 0 & \frac{3}{32} & 0 & \frac{1}{16} & 0 & \frac{1}{64} \\ \frac{1}{256} & 0 & \frac{1}{64} & 0 & \frac{3}{128} & 0 & \frac{1}{64} & 0 & \frac{1}{256} \end{pmatrix} \quad (4)$$

2.2.2. Salient feature based on cluster analysis

Here, we adopt the mean-shift cluster algorithm [21] to analyze image content from the wide-field image. Mean-shift is a data clustering algorithm commonly used in image segmentation. Compared with other popular clustering methods such as the k-means method [22], mean-shift does not need to define the number of clusters. In mean-shift, it considers the pixels of the input image as sampled from the underlying probability density function and constantly locating the maximum of a density function. In wide-field, considering there is a significant difference in the intensity value between the biological sample structure and background. We adopt a mean-shift to distinguish the wide-field image content. With the advanced defined hyper-parameters, spatial radius r and intensity feature distance d , mean-shift replaces each pixel with the mean of the pixels in a range r neighborhood and whose intensity is within distance d . At next iteration, mean-shift calculate a shift vector $Ms(x_i^t)$ (Eq. (5)) to move the region to the location of the new centroid. When the variation of $Ms(x_i^t)$ in the last two iterations is less than a threshold or the iteration meets a certain number, it indicates that the mean-shift vector has converged. Thus, pixels with the same or similar intensity will be assigned the same category labels, thereby achieving salient structures detection.

$$Ms(x_i^t) = \frac{\sum_{j=i-r}^{i+r} x_j g\left(\left\|\frac{x_i - x_j}{d}\right\|\right)}{\sum_{j=i-r}^{i+r} g\left(\left\|\frac{x_i - x_j}{d}\right\|\right)} - x_i \quad (5)$$

$$g\left(\left\|\frac{x_i - x_j}{d}\right\|\right) = \begin{cases} \exp\left(-\frac{1}{2} \left\|\frac{x_i - x_j}{d}\right\|^2\right), & \text{if } \left\|\frac{x_i - x_j}{d}\right\| \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

The spatial radius r and intensity feature distance d are the only parameters involved in DETECTOR. Here, these two parameters determine the sensitivity to clusters. Generally, small spatial radius and small intensity features result in high sensitivity. While large spatial radius and large intensity features result in low sensitivity. The higher the clustering sensitivity, the greater the number of small clusters. Thereby, the clustering results with small parameters can retain more image details. Here, Fig. 2 shows cluster results with different parameters. In detail, Fig. 2(B) is set with $r = 5$ and $d = 5$. Figure 2(C) is set with $r = 15$ and $d = 15$. From the results, Fig. 2(B) contains more detail structures than Fig. 2(C). Considering that the salient feature is used in the weight mask computation, to obtain a more representative weight value according to the image content, we recommend the use of small r and d in the mean-shift method.

2.3. Super-resolution image degradation model

2.3.1. Pyramid sub-pixel registration method

The alignment between the super-resolution image and the wide-field image is necessary, especially for application of dual channel imaging or the imaging detector across different areas. Here, we adopt a classic pyramid sub-pixel registration method [23] to align the super-resolution image to the wide-field image. We first establish a pyramid image for each reference and test data. The pyramid images offer a multi-scale resolution presentation of the super-resolution image and the wide-field image by down-sampling. First, the largest-scale resolution images achieve the initial alignment with the minimum details. Then we used a coarse-to-fine strategy. The iterative corrections are made for finer details in pyramid image data. Finally, we can get an excellent aligned super-resolution image. By comparing the merged image after alignment with the previous merged image, we could see the success of alignment (Fig. 3).

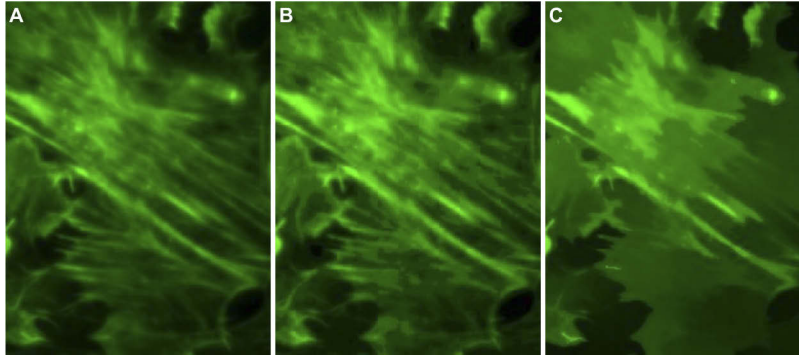


Fig. 2. Salient features of the wide-field image by mean-shift clustering. (A) The wide-field images. (B) The clustering results of the wide-field image by a mean-shift algorithm with r as 5 and d as 5. (C) The clustering results of the wide-field image by a mean-shift algorithm with r as 10 and d as 10.

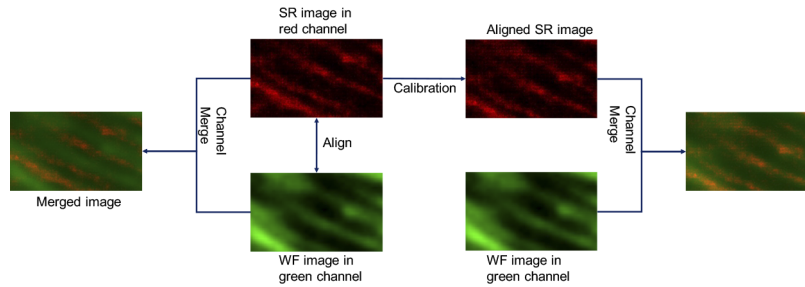


Fig. 3. The workflow of alignment. This figure contains two merging processes. The left subfigure is before alignment, and the right subfigure is after alignment.

2.3.2. PSF Gaussian model

When alignment is completed, DETECTOR rescales the resolution of the aligned super-resolution image with the PSF model measured from the optical imaging system. Generally, PSF is mathematically modeled as a Gaussian function at the point (x,y) (Eq. (7)) ([12,13,24]). Measuring the full width at half maximum (FWHM) is the most practical way to characterize PSF. The FWHM of the curve function is the distance between the points where the intensity is half of the maximum one. To measure FWHM, a 3D image stack with 50 fluorescent beads is obtained at different Z planes. Then we computed the FWHM of each fluorescent bead according to their intensity curve and obtained lateral FWHM statistical distribution [25]. The relationship with parameter σ of PSF gaussian model between FWHM is formulated by Eq. (8). DETECTOR uses the mean value of FWHM statistical distribution to compute σ value and gets the actual PSF model.

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (7)$$

$$FWHM = 2\sqrt{2 \ln 2} \sigma \approx 2.355\sigma \quad (8)$$

2.4. Super-resolution structural similarity metric

To detect artifacts of super-resolution image, besides introduce a weight mask, here we propose a novel structural similarity metric based on SSIM. The SSIM assesses the image based on the assumption that human visual perception (HVS) is highly adapted for extracting structural

information from a scene. SSIM (Eq. (9)) assesses the similarity of two tested image m, n based on luminance, contrast and structure comparison. These three component comparisons are separately noted as $l(m, n)$, $c(m, n)$ and $s(m, n)$. The detailed computation of each above component is shown in Eq. (10), Eq. (11) and Eq. (12). In Eq. (9), the α , β and γ are the computing weights of the above three components. The constant term $C_i (i = 1, 2, 3)$ in denominator of Eq. (10), Eq. (11) and Eq. (12) is to avoid that the denominator tends to zero. The formula of C_1 and C_2 are shown in Eq. (13) and Eq. (14). The L means the dynamic range of the pixel values. Here, with reference to the parameters setting in SSIM paper [26], $k_1 = 0.01$ and $k_2 = 0.03$; $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$. Thus, Eq. (9) is can be inferred to the form of Eq. (15).

However, due to the intensity gap of degraded super-resolution images and wide-field images. Thus, it is prone to report artifact which does not exist if directly applying the original SSIM metric because SSIM metric contains the luminance comparison. Here, to weaken the influence of intensity gap, DETECTOR proposed a novel similarity metric, named MASK-SSIM. MASK-SSIM metric derives from the SSIM metric. Different from SSIM metric, first, considering intensity comparison has no significant contribution to the original SSIM metric [27], MASK-SSIM metric only preserves contrast and structural comparison. Second, to focus on regions where biological structures exist, MASK-SSIM contains a weight mask that is extracted from the wide-field image to enhance the structural similarity. The formula of MASK-SSIM is shown in Eq. (16). The output value of the MASK-SSIM metric ranges from 0 to 1 which is positively correlated with the image similarity. Specifically, MASK-SSIM metric outputs 1 when input two identical images and outputs 0 when two test images are entirely different.

Based on the sliding window, DETECTOR outputs error map and similarity value by computing the MASK-SSIM of degraded super-resolution image and wide-field image. All error maps of DETECTOR are from 0-255. This is due to the value range of the MASK-SSIM formula. In the visualization of error maps, to highlight artifacts of super-resolution images, we design an error map where higher intensity refers to less convince of the reconstructed images. Thus, for the 8-bit error map image, the relationship between the pixel intensity with MASK-SSIM is shown in Eq. (17). The $S_{(x,y)}$ refers to the MASK-SSIM value at location (x,y) within the sliding window and the $I_{(x,y)}$ refers to the intensity of the error map at location (x,y) . Here, we recommend adopting the sliding window with size 3 (Supplement 5). The error map of DETECTOR is of the same size as the super-resolution image, where the intensity value of each pixel presents the MASK-SSIM value of the sliding window. The overall similarity value is the average value of the error map pixels.

$$SSIM(m, n) = [l(m, n)^\alpha \cdot c(m, n)^\beta \cdot s(m, n)^\gamma] \quad (9)$$

$$l(m, n) = \frac{2\mu_m\mu_n + C_1}{\mu_m^2 + \mu_n^2 + C_1} \quad (10)$$

$$c(m, n) = \frac{2\sigma_m\sigma_n + C_3}{\sigma_m^2 + \sigma_n^2 + C_3} \quad (11)$$

$$s(m, n) = \frac{\sigma_{mn} + C_2}{\sigma_m + \sigma_n + C_2} \quad (12)$$

$$C_1 = (k_1L)^2 \quad (13)$$

$$C_2 = (k_2L)^2 \quad (14)$$

$$SSIM'(m, n) = \frac{(2\mu_m\mu_n + C_1)(2\sigma_{mn} + C_2)}{(\mu_m^2 + \mu_n^2 + C_1)(\sigma_m^2 + \sigma_n^2 + C_2)} \quad (15)$$

$$MASK - SSIM(m, n) = \sum_{i,j \in \Omega} wc(i, j)s(i, j) \quad (16)$$

where Ω means the sliding window subregion; w means the weight value of the corresponding sliding region, μ_m and μ_n are the intensity of image m and image n ; σ_m^2 and σ_n^2 are the variance

of image m and image n ; σ_m and σ_n are the covariances of image m and image n ; σ_{mn} is the covariance of image m and image n .

$$I_{(x,y)} = (1 - S_{(x,y)}) * 255 \quad (17)$$

3. Results

3.1. Data preparation

Two simulated datasets and three experiential datasets are used to evaluate the performance of DETECTOR. The simulated data are microtubule data and grid data. The microtubule dataset (Fig. 4) is a public synthetic dataset designed in a competition that aims to rank the performance of 2-dimension SMLM software packages [8]. It shows a realistic microtubule structure with seven thin tubule structures (constant diameter 25 nm) and one thick tubule structure (constant diameter 40 nm). In order to design the most realistic synthetic data, the designers consider experimentally background and signal-to-noise levels based closely on common experimental conditions. In the dataset, there are three independent sources of photons: the signal of interest (activated molecule), the background signal normally distributed, which slowly changes with time, and the autofluorescent signal simulated by introducing deep clusters of intense fluorophores that are constantly in an active mode, slowly change with time. The grid data was synthesized by ThunderStorm (Fig. 5) [28]. The grid data comprises six intersecting lines with gradually decaying brightness from top to bottom and left to right. The grid data set consisted of only 3000 image frames with low molecular density: $3\mu m^2$. The super-resolution image is a direct display of the single molecules overall with their localization information. The image size is 60×60 pixels (1 pixel = 100 nm). Here, we set the FWHM of the PSF model as 210 nm and the range of the total number of photons from 100 to 600.

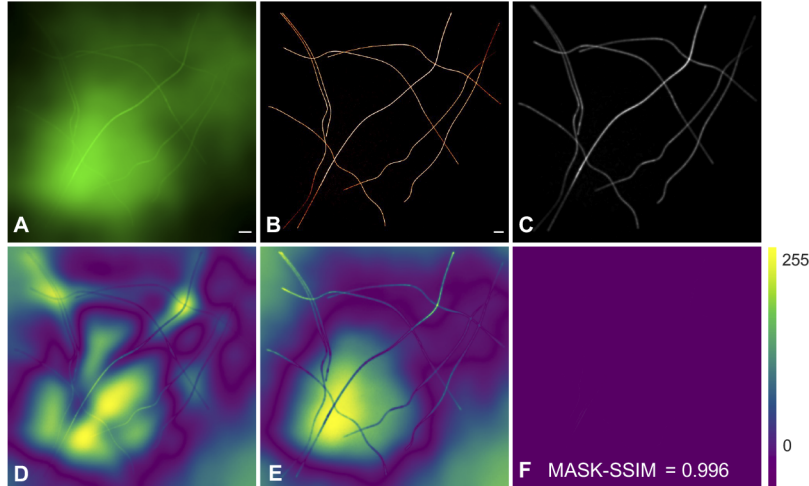


Fig. 4. The performance of DETECTOR on simulated public data with strong autofluorescent background. (A) The wide-field image; (B) The simulated STORM super-resolution image with actual single molecules positions; (C) The degraded super-resolution image obtained by actual PSF. (D) The SQUIRREL error map of wide-field image and the degraded super-resolution image obtained by RSF. (E) The SQUIRREL error map of wide-field image and the degraded super-resolution image obtained by actual PSF. (F) The error map of DETECTOR. The color bar is for D, E and F. In the error map, a higher intensity value means that there is more dissimilarity between wide-field image and super-resolution image. The scale bar is $10\mu m$.

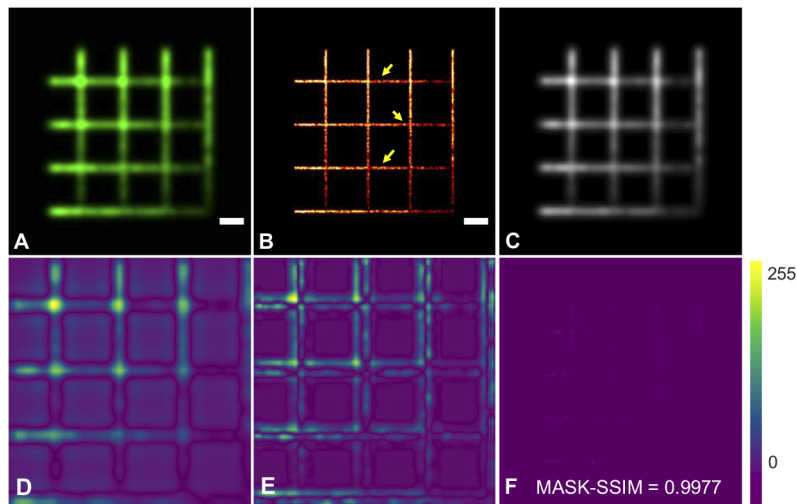


Fig. 5. The performance of DETECTOR on super-resolution image of SMLM. (A) The wide-field image (B) The simulated STORM super-resolution image with actual single-molecule positions. (C) The degraded super-resolution image obtained by actual PSF model. (D) The error map of SQUIRREL with RSF. (E) The error map of SQUIRREL with actual PSF model. (F) The error map of DETECTOR. The scale bar is $1 \mu\text{m}$. As SQUIRREL cuts the border regions of input images, the size of A-C and D-E is different.

All experimental data sets are image sequences of cellular structures labeled by the photo-convertible fluorescent protein (PCFP) mEos3.2. This protein contains two imaging modalities. Under a 405 nm laser, mEos3.2 can convert from green color to red color. In the red channel, isolated fluorophores can be precisely localized by the PALM. In the green channel, the high-density fluorophores perform very well with on/off or blinking and bleaching phenomena, which provides the fluctuation information of the image sequence in the time domain. The cellular structures of these experimental data include two actin networks in U2OS cells named actin1 (Fig. 7), actin2 (Fig. 6). The clathrin-coated pits (CCP) structure in HeLa cells named CCP data (Fig. 8).

For acquisition, we used a custom-built total internal reflection fluorescence (TIRF) microscopy system with an Olympus IX71 body (Olympus), high-NA oil objectives, and an electron-multiplying charge-coupled device (EMCCD) camera (Andor iXon DV-897 BV). For actin data, the image pixel size of 160 nm was determined by a $100\times$, 1.49 NA oil objective (Olympus PLAN APO). For the CCP data, the image pixel size of 66.7 nm was determined by a $150\times$, 1.45 NA objective, and $1.6\times$ intermediate magnifications.

Therefore, each experimental dataset consists of two sets of image sequences, one is from the red channel and the other is from the green channel. For actin1, we obtained 20,000 frames of red channel data and 200 frames of green channel data. For actin2, we obtained 50,000 frames of red channel data and 200 frames of green channel data. For CCP data, we obtained 5,000 frames of red channel data and 200 frames of green channel data. To obtain the super-resolution images, for frames collected by Photo-activated localization microscopy (PALM) [13], we adopted Gaussian fitting to locate individual molecules in each image frame and obtained the super-resolution image. Here, the super-resolution image labeled with PALM means the frames collected by PALM and reconstructed by Gaussian fitting. For frames of green channel data which contain high-density molecules, we adopted Super-resolution radial fluctuations (SRRF) [29] and Single molecule-guided Bayesian localization microscopy (SIMBA) [30,31] to obtain the super-resolution image. Specifically, we adopted all 200 frames of the green channel signal to reconstruct the SRRF image. For SIMBA, we used the first 200 red

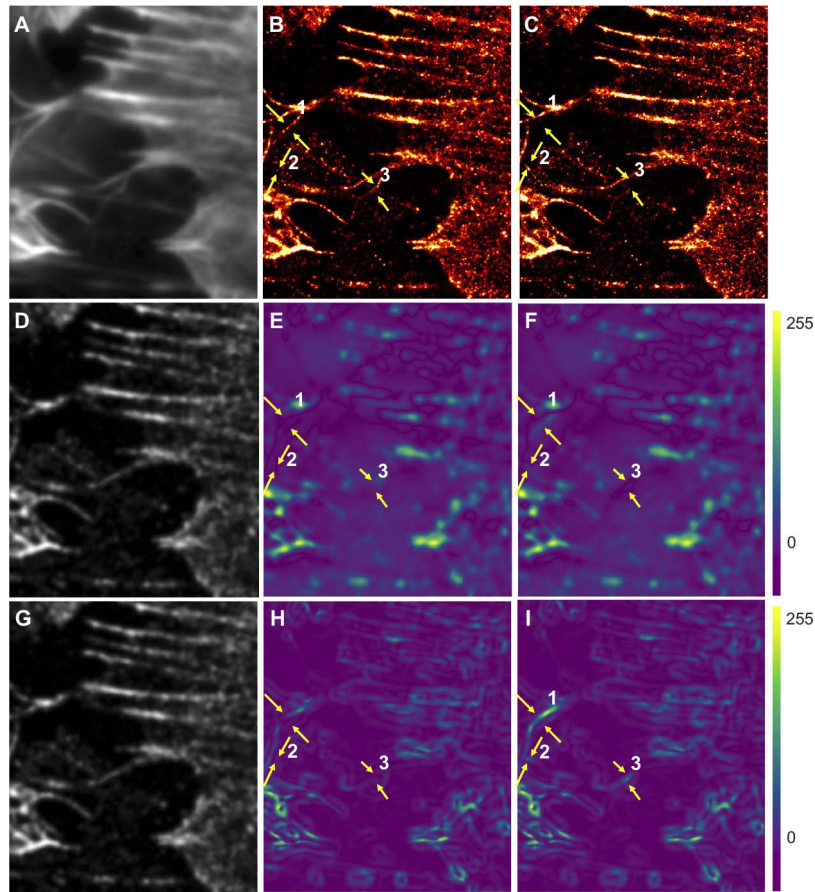


Fig. 6. The artifact detection performance of DETECTOR. (A) The wide-field image. (B) The super-resolution image reconstructed by PALM. (C) The modified PALM super-resolution image with removed and added structures (arrows). (D) The degraded super-resolution image generated with RSF. (E)(F) The error map of the original and modified PALM super-resolution image by SQUIRREL individually. (G) The degraded super-resolution image generated with actual PSF. (H)(I) The error map of the original and modified PALM super-resolution image by DETECTOR.

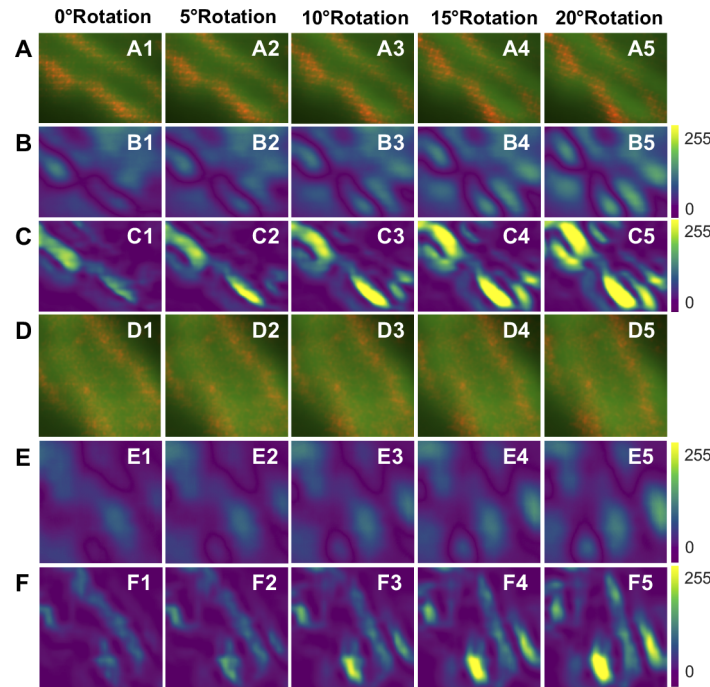


Fig. 7. The sensitivity to structural distortions. Row A-C and D-F represent two subregion examples. Each column corresponds to one rotation angle, from left to right: 0° rotation, 5° rotation, 10° rotation, 15° rotation and 20° rotation. For each example, the first row (A, D) presents the merged image where the wide-field image is colored in green and the rotated degraded super-resolution image is colored in red. The second row (B, E) presents the error map of SQUIRREL. The third row (C, F) presents the error map of DETECTOR.

channel frames for single-molecule position extraction then used all 200 green channel frames for Bayesian calculation to get the final super-resolution image.

3.2. Performance on the simulated dataset

3.2.1. DETECTOR is robust to autofluorescent background

In this experiment, we tested DETECTOR on a public synthetic dataset to show it only focus on artifacts of biological structures in super-resolution images. In this public synthetic dataset, there is a wide-field image (Fig. 4(B)) that contains strong autofluorescent background and a ground-truth super-resolution image (Fig. 4(A)) which accurately shows structures of the wide-field image. This means if we rescale the resolution of the super-resolution image to be the same as that of the wide-field image, the structural information of these two images is the same. Figure 4(C) shows the resolution rescaled super-resolution image. Then, we computed the similarity between the degraded super-resolution image and wide-field image in DETECTOR and obtained an error map and MASK-SSIM value (Fig. 4(F)). In the error map, a higher intensity value means that there are more dissimilarities between wide-field image and super-resolution image. As DETECTOR only focuses on artifacts of biological structures, this error map does not show any obvious highlight regions. This makes sense because the only difference between a wide-field image and a degraded super-resolution image is caused by the strong autofluorescent background and noise.

Also, we tested SQUIRREL on this public synthetic dataset. When rescaling the resolution of the super-resolution image, we tested two ways which are supported by SQUIRREL. The detail of

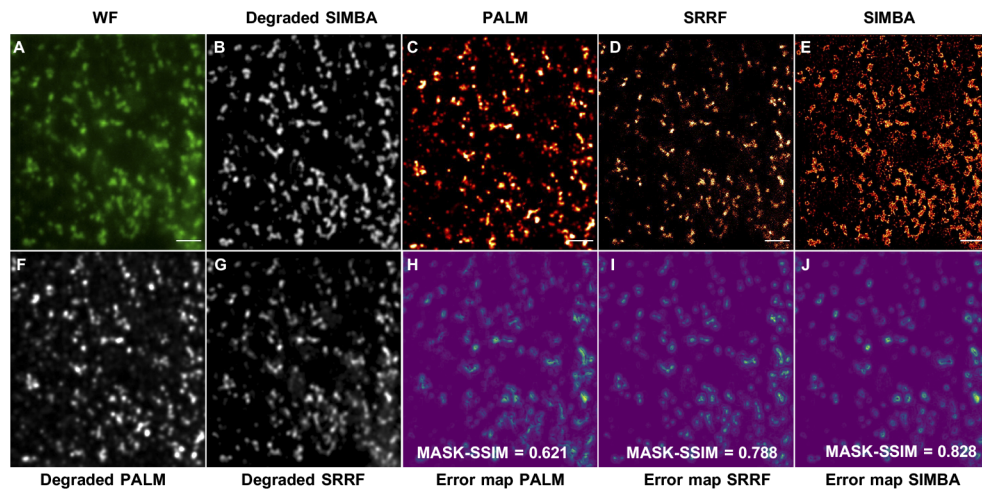


Fig. 8. Quality evaluation of different super-resolution images reconstructed by SRRF, SIMBA and PALM with CCP data. (A) The wide-field image. (B) The degraded super-resolution image of SIMBA. (C) The PALM super-resolution image. (D) The SRRF super-resolution image. (E) The SIMBA super-resolution image. (F) The degraded super-resolution image of PALM. (G) The degraded super-resolution image of SRRF. (H) The error map corresponding to the PALM super-resolution image. (I) The error map corresponding to the SRRF super-resolution image. (J) The error map corresponding to the SIMBA super-resolution image. The indicated score is the MASK-SSIM metric value. The scale bar is 10 μm .

the degraded super-resolution image of SQUIRREL can see the Fig. S1. For further discussion, here we briefly describe these two degraded methods. The one is the same as DETECTOR, SQUIRREL supports rescaling resolution with actual PSF that is modeled by the Gaussian model. The other is that SQUIRREL estimates a resolution scaling function (RSF) by optimization. Figure 4(D) shows the error map of the wide-field image and the degraded super-resolution image obtained by RSF. Figure 4(E) shows the error map of the wide-field image and the degraded super-resolution image obtained by actual PSF.

In the error map of Fig. 4(D) and Fig. 4(E). It makes sense that SQUIRREL shows highlight regions in non-structures regions because it detects artifacts by computing intensity mismatch. Whether using the actual PSF or the RSF, autofluorescent background can affect the results. Thus, for the dataset with strong autofluorescent background, if users focus more on the autofluorescent background, SQUIRREL is more suitable. While, if users want to reduce the effect of autofluorescent background on the results DETECTOR is more suitable.

3.2.2. DETECTOR is robust to super-resolution images of SMLM

To prove that DETECTOR is robust to super-resolution images of SMLM, we tested DETECTOR on a synthetic dataset whose intensity of lines gradually changes. Figure 5(A) shows the wide-field image of this dataset where intensity value of each line gradually decreases from left to right and from top to bottom. Figure 5(B) is the ground truth image provided by ThunderSTORM simulation method. ThunderSTORM is one of the SMLM methods. Here, we designed a synthetic dataset with gradually decreased intensity to show the molecules whose intensity has sudden changes. This intensity sudden change is not a labeling density artifact that presents in both the wide-field image and super-resolution image but is caused by an intensity information mismatch between these two images. For a wide-field image, the intensity of each pixel depends on the label density of fluorescent molecules and the intensity of activated molecules. While, for

super-resolution images reconstructed by SMLM, the intensity of each pixel is depend on the label density of fluorescent molecules and the blinking number of every single molecule. This intensity information mismatch of the two images can lead to bias when calculating the error map depending on the intensity.

It is worth emphasizing that this artifact refers to the molecules whose brightness suddenly changes in the reconstructed SMLM images rather the intensity gradually decrease of line structures. Similar to the above experiment, we both tested SQUIRREL and DETECTOR on the grid data. The error maps are shown in Fig. 5(D)–5(F). Here, Fig. 5(D) and 5(E) are the error maps of SQUIRREL. One is computed with the RSF and the other is computed with the actual PSF model. The detail of the degraded super-resolution image of SQUIRREL can see the Fig. S2. Here, Fig. 5(F) is the error map of DETECTOR. From the results, we can see that although SQUIRREL linearly adjusts the intensity of the super-resolution image to match the intensity of wide-field image maximally. Figure 5(D) and Fig. 5(E) shows highlight areas where there exist an intensity gap. As DETECTOR only focus on structural similarity, the error map of DETECTOR contains no highlight areas and the MASK-SSIM value is close to 1.

3.3. Performance with experimental dataset

3.3.1. Artifact detection performance of DETECTOR

In this experiment, we selected a subregion of actin1 data to evaluate the artifact detection performance of DETECTOR (Fig. 6). The wide-field image of actin1 is shown in Fig. 6(A) and the corresponding super-resolution image which PALM reconstructs is shown in Fig. 6(B). To obtain a super-resolution with structural artifacts (Fig. 6(C)), we deliberately removed two filaments (arrows marked as 1 and 3) and added one filament (arrow marked as 2) from the raw super-resolution image. We tested the raw super-resolution image and the modified super-resolution image on both SQUIRREL and DETECTOR. With the measured real PSF model, SQUIRREL and DETECTOR separately rescale the resolution of the super-resolution image (Fig. 6(D),G) and output corresponding error maps. The output of error maps of SQUIRREL are shown in Fig. 6(E) and Fig. 6(F). The output of error maps of DETECTOR are shown in Fig. 6(H) and Fig. 6(I). From the error map, we can see that both DETECTOR and SQUIRREL can clearly detect the artifacts in the subregion 1. For the other missing substructure (arrow marked as 3), only DETECTOR shows corresponding highlighted regions on the error map. For the additional substructure (arrow marked as 2), DETECTOR also highlights this artifact. The results demonstrate the capability and advantage of DETECTOR to identify artifacts even at low signal-to-noise images accurately. Besides, we tested both DETECTOR and SQUIRREL on a public dataset [11] to show more convince results (Supplement 6).

3.3.2. DETECTOR is sensitive to small structural distortions

Besides identifying artifacts of images with low signal-to-noise, DETECTOR can discriminate slight structural distortions. We chose two subregions of actin2 data and rotated the reconstructed structures of the super-resolution image with different degrees: 0° , 5° , 10° , 15° , 20° . To visualize the structural deformation, we color the wide-field image in green, color the corresponding degraded super-resolution image in red, and then merge the two images (Fig. 7(A, D)). From Fig. 7(A, D), we can see that as the rotation angle increased, the mismatches between the degraded super-resolution image and the wide-field images became larger and larger. We tested raw super-resolution images and all rotated images both on SQUIRREL and DETECTOR. Figure 7(B),C are the results of SQUIRREL and DETECTOR on subregion 1. Figure 7(E),F are the results of SQUIRREL and DETECTOR on subregion 2. From the results, we can see that, both SQUIRREL and DETECTOR can detect structural distortions. When the rotation of structures is 0° , there are still highlight areas in error maps. This is because that the degraded super-resolution image is not exactly the same as the wide-field image. We convolved the

super-resolution image with actual PSF to achieve the resolution consistency of the degraded super-resolution image with the wide-field image. At this resolution level, we detect artifacts by image similarity assessment. From Fig. 6(B1) and Fig. 6(C1), both SQUIRREL and DETECTOR show highlight areas. Thus, we can conclude that the super-resolution image may exist some artifacts. Besides, in order to make the experiment more convince and equivalent, we tested a dataset from SQUIRREL (Supplement 7). As the biological sample becomes more distorted, the highlight regions of these two kinds of error maps are more bright. Particularly, even when the rotation angle is 5° , DETECTOR highlight the structural distortions of the corresponding error map.

3.3.3. Guiding the selection of super-resolution techniques

Besides detecting structural artifacts, DETECTOR can help guide choosing the ideal super-resolution reconstruction model. Here, we used three kinds of reconstruction methods to obtain a super-resolution image of CCP data. The wide-field image of CCP data is shown in Fig. 8(A). The three reconstruction methods are PALM, SRRF and SIMBA, and their corresponding super-resolution image is shown in Fig. 8(C), 8(D) and 8(E). From these super-resolution images, we can see that except for PALM, both SRRF and SIMBA can reconstruct the ring structure of CCP. Furthermore, SIMBA is more suitable for this data because it reconstructs more fine structures. Theoretically, with a sufficient frame series, PALM can present the highest resolution image of this data. However, because this dataset only collects 5,000 original frames, it is reasonable that the resolution of PALM reconstructed images is worse than SRRF and SIMBA. Figure 8(B), 8(F), 8(G) show the degraded super-resolution image, which is obtained with a real measured PSF model. Figure 8(H), I, J shows the error maps of DETECTOR and their corresponding MASK-SSIM values. According to the results, we can find that the value of MASK-SSIM is positively correlated with the quality of the reconstructed images. Thus, for this CCP data, SIMBA can reconstruct a high-quality super-resolution image.

3.3.4. Guiding parameter tuning of the reconstruction model

For some reconstruction methods (such as SIMBA), key parameters may affect the quality of the reconstructed super-resolution image. Therefore, how to quickly select the optimal parameters is an important issue. With the guidance of MASK-SSIM, DETECTOR can help users tuning parameters during reconstruction. In SIMBA, the low-pass filter (lpf) is a key parameter in reconstruction, which affects the analysis of low-frequency and high-frequency information in the frame series. The lpf value can be set as: 1, 3, and 5. Higher lpf value can lead to smoother structures of super-resolution images. Here we use SIMBA to reconstruct CCP super-resolution images, where lpf is set to 1, 3 and 5. As shown in Fig. 9, Fig. 9(A) is the wide-field image of CCP data and the three super-resolution images are shown in Fig. 9(B)-D. Here, we tested these three images on DETECTOR and obtained the corresponding MASK-SSIM values. The results were quite consistent with a priori knowledge. When the lpf was set to 1, the low-frequency information was lost, leading to incomplete structures and the lowest quality score in the super-resolution image. The situation was improved when the ring structures began to emerge when the lpf was set to 3. When increasing lpf to 5, the super-resolution image presents a smoother structure and sacrifices the sharp edges. The quality score of SIMBA super-resolution images with lpf set to 3 and 5 are similar values. However, Fig. 9(C) contains much more sharp edges, leading to the highest quality score among the scores with the three settings. The results demonstrated that DETECTOR helps tune parameters in super-resolution image reconstruction.

3.3.5. Guiding data collection of the PALM

Generally, in the super-resolution reconstruction method, the number of frames during data collection is positively correlated with the resolution of the reconstructed image. However, collecting

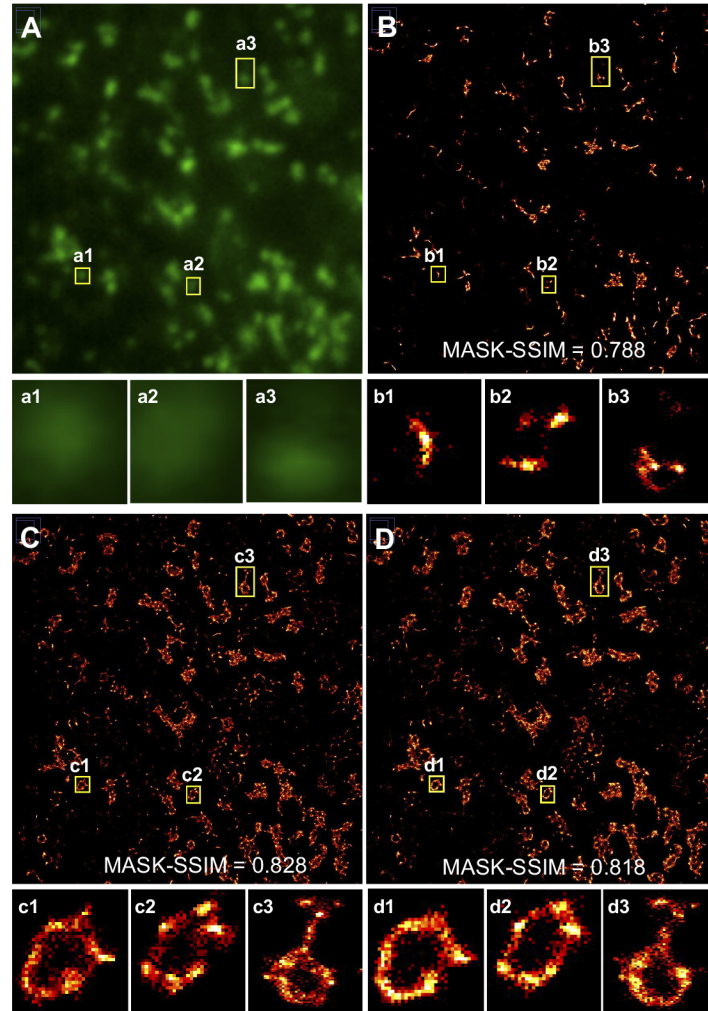


Fig. 9. Guiding parameter tuning of the reconstruction model. (A) The wide-field image of the CCP data. (B)-(D) The SIMBA super-resolution images with $lpf = 1, 3$ and 5 . Magnified images of subregions (yellow box) are presented below each image (A-D). The scale bar is $10 \mu m$.

longer frames requires more acquisition time and can lead to phototoxicity, photobleaching and sample drift. Therefore, it is crucial to choose a sufficient number of frames during data collection. Besides guiding parameters tuning of the reconstruction model, DETECTOR can also guide the data collection. Here, we collect 10 sets of actin2 data. The number of frames of each set is (5,000, 10,000, 15,000 \cdots 50,000) and each set is increased by 5,000 frames. With these datasets, we reconstructed ten PALM super-resolution images. Figure 10 shows a sub-region of the wide-field image of actin2 data and its corresponding ten reconstructed images. From the figures, we can see that as the number of frames increases, the super-resolution image presents more fine detail structures. Then we tested these ten set images on DETECTOR and analyzed the MASK-SSIM values of each dataset (the bottom right figure of Fig. 10). In Fig. 10, we can see that the MASK-SSIM value gradually increases as the number of frames increase until it reached 20,000. After 20,000, the MASK-SSIM value remains stable. Thus, we can conclude that collecting 20,000 frames would be sufficient for PALM reconstruction. With the guidance of DETECTOR, we can choose a suitable number of frames during the data collection to avoid sample drift and damage.

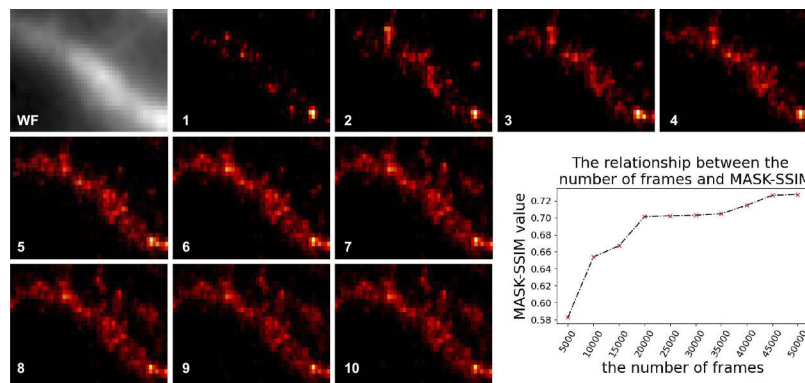


Fig. 10. Ten super-resolution images reconstructed by PALM and their MASK-SSIM value on DETECTOR. The presented images are one subregion of the whole actin2 image. Label 1 indicates the super-resolution image reconstructed from 5,000 frames; Label 2 indicates the super-resolution image reconstructed from 10,000 frames and so on. The bottom right figure is a line chart that presents the relationship between the number of frames and the MASK-SSIM index. The abscissa represents the number of image frames, and the unit is 5,000 frames. The ordinate is MASK-SSIM value.

4. Discussion and conclusion

In this work, we proposed an image artifact detection method (DETECTOR) for super-resolution. Different from the existing methods, our method focus on structural artifacts. Our method contains three key features to make the image similarity assessment focusing on structural information. We used the actual PSF measured from the optical imaging system to make the resolution consistency between the wide-field image and the degraded super-resolution image. We introduced a weight mask extracted to filter out the strong autofluorescence background and proposed a novel MASK-SSIM index for super-resolution image assessment.

The simulated experimental results supported that DETECTOR has an advantage in assessing images of biologically thick samples with strong autofluorescence background, such as tissue slices, embryos and tiny organisms. The simulated experimental results also exhibited that DETECTOR is not susceptible to distortion caused by sudden changes in the molecular intensity

of the fluorescence image reconstructed by SMLM. Moreover, DETECTOR is surprisingly sensitive to minor dissimilarities, including distortions as small as 5° .

To verify the reliability of our method for assessing the quality of real experimental data, especially those reconstructed by SMLM imaging methods, we manually removed and added substructures in the PALM image. According to the error map of these super-resolution images, our method can clearly identify artifacts distributed on the weak signal region. We evaluated super-resolution images from SRRF, SIMBA and PALM/STORM and super-resolution images reconstructed with different parameter settings. The experimental results show that DETECTOR can offer valuable guidance for users to select the most appropriate imaging method or the optimal reconstruction parameter for a given data. We also demonstrated that DETECTOR can guide data imaging strategies, such as defining a criterion for acquiring a sufficient amount of data. We envision that DETECTOR would help choose fluorescence labeling strategies as well.

In summary, our study provides a new computation method for artifact detection in super-resolution images. Compared with the state-of-the-art method, DETECTOR focus on structural similarity between the wide-field image and degraded super-resolution image. Thereby, it can avoid a false alert artifact that is caused by the intensity mismatch between these two images. For example, wide-field images of biologically thick samples which contain strong autofluorescent background and SMLM reconstructed images whose intensity information is inconsistent with that of wide-field images. With the experiment of structures rotation, we also found that DETECTOR is highly sensitive to structural distortion artifacts. Besides, with the guidance of MASK-SSIM, DETECTOR can help users to select the most appropriate imaging method or the optimal reconstruction parameter for a given data.

However, DETECTOR has some limitations due to the use of wide-field images as a reference. First, the degradation of super-resolution images eliminates fine structures in the image, thereby limiting the precision of detecting artifacts. Second, a wide-field image with high SNR is essential to ensure the reliability of DETECTOR. We tested DETECTOR on wide-field images with different SNR. We found that the structures of the wide-field images can be damaged when the image contains high-level noise, which can affect the similarity computation between the wide-field image and the degraded super-resolution image.

Funding. National Key Research and Development Program of China (2017YFA0504702, 2017YFA0505300); National Natural Science Foundation of China (21778069, 21927813, 31870857, 32027901, 61932018, 62072441); Beijing Municipal Natural Science Foundation (L182053); The Strategic Priority Research Program of Chinese Academy of Sciences (XDB37040301).

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request. The source code of DETECTOR are freely available in Ref. [32].

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. Y. M. Sigal, R. Zhou, and X. Zhuang, "Visualizing and discovering cellular structures with super-resolution microscopy," *Science* **361**(6405), 880–887 (2018).
2. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, and X. Zhuang, "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging," *Nat. Methods* **8**(12), 1027–1036 (2011).
3. P. Fox-Roberts, R. Marsh, K. Pfisterer, A. Jayo, M. Parsons, and S. Cox, "Local dimensionality determines imaging speed in localization microscopy," *Nat. Commun.* **8**(1), 13558–10 (2017).
4. A. Burgert, S. Letschert, S. Doose, and M. Sauer, "Artifacts in single-molecule localization microscopy," *Histochem. Cell Biol.* **144**(2), 123–131 (2015).
5. R. J. Marsh, K. Pfisterer, P. Bennett, L. M. Hirvonen, M. Gautel, G. E. Jones, and S. Cox, "Artifact-free high-density localization microscopy analysis," *Nat. Methods* **15**(9), 689–692 (2018).
6. P. Almada, S. Culley, and R. Henriques, "Palm and storm: Into large fields and high-throughput microscopy with sncmos detectors," *Methods* **88**, 109–121 (2015).
7. J. Sinkó, R. Kákonyi, E. Rees, D. Metcalf, A. E. Knight, C. F. Kaminski, G. Szabó, and M. Erdélyi, "Teststorm: Simulator for optimizing sample labeling and image acquisition in localization based super-resolution microscopy," *Biomed. Opt. Express* **5**(3), 778–787 (2014).

8. D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "quantitative evaluation of software packages for single-molecule localization microscopy," *Nat. Methods* **12**(8), 717–724 (2015).
9. C. S. Smith, J. A. Slotman, L. Schermelleh, N. Chakrova, S. Hari, Y. Vos, C. W. Hagen, M. Müller, W. van Cappellen, A. B. Houtsmuller, J. P. Hoogenboom, and S. Stallinga, "Structured illumination microscopy with noise-controlled image reconstructions," *Nat. Methods* **18**(7), 821–828 (2021).
10. J. J. Bozzola and L. D. Russell, *Electron Microscopy: Principles and Techniques for Biologists* (Jones & Bartlett Learning, 1999).
11. S. Culley, D. Albrecht, C. Jacobs, P. M. Pereira, C. Leterrier, J. Mercer, and R. Henriques, "Quantitative mapping and minimization of super-resolution optical imaging artifacts," *Nat. Methods* **15**(4), 263–266 (2018).
12. B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.* **46**(10), 1819–1829 (2007).
13. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642–1645 (2006).
14. S. K. Maji and H. Yahia, "A feature based reconstruction model for fluorescence microscopy image denoising," *Sci. Rep.* **9**(1), 7725–7729 (2019).
15. G. Liu, F. Liu, and A. T. Usman, "Image feature extraction under microscope based on visual saliency," *Acta Microsc.* **29**(4), 1892–1899 (2020).
16. R. Arulmurugan and H. Anandakumar, "Region-based seed point cell segmentation and detection for biomedical image analysis," *Int. J. Biomed. Eng. Technol.* **27**(4), 273–289 (2018).
17. J. Li, Z. Pan, Q. Liu, Y. Cui, and Y. Sun, "Complementarity-aware attention network for salient object detection," *IEEE transactions on cybernetics* (2020).
18. R. Muthukrishnan and M. Radha, "Edge detection techniques for image segmentation," *Int. J. Comput. Sci. & Inf. Technol.* **3**(6), 259–267 (2011).
19. P. Dutilleul, "An implementation of the "algorithme à trous" to compute the wavelet transform," in *Wavelets* (Springer, 1990), pp. 298–304.
20. W. Rakowski, "Prefiltering in wavelet analysis applying cubic b-splines," *Int. J. Electron. Telecommun.* **60**(4), 331–340 (2014).
21. K. G. Derpanis, "Mean shift clustering," *Lect. Notes* p. 32 (2005).
22. J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *J. royal statistical society. series c (applied statistics)* **28**(1), 100–108 (1979).
23. P. Thevenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. on Image Process.* **7**(1), 27–41 (1998).
24. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm)," *Nat. Methods* **3**(10), 793–796 (2006).
25. C. Matthews and F. Cordelieres, "Metroloj: an imagej plugin to help monitor microscopes health," in *ImageJ User & Developer Conference proceedings*, (2010).
26. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
27. D. M. Rouse and S. S. Hemami, "Understanding and simplifying the structural similarity metric," in *2008 15th IEEE International Conference on Image Processing*, (IEEE, 2008), pp. 1188–1191.
28. M. Ovesný, P. Krížek, J. Borkovec, Z. Švindrych, and G. M. Hagen, "Thunderstorm: a comprehensive imagej plug-in for palm and storm data analysis and super-resolution imaging," *Bioinformatics* **30**(16), 2389–2390 (2014).
29. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, "Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations," *Nat. Commun.* **7**(1), 12471 (2016).
30. F. Xu, M. Zhang, W. He, R. Han, F. Xue, Z. Liu, F. Zhang, J. Lippincott-Schwartz, and P. Xu, "Live cell single molecule-guided bayesian localization super resolution microscopy," *Cell Res.* **27**(5), 713–716 (2017).
31. H. Li, F. Xu, S. Gao, M. Zhang, F. Xue, P. Xu, and F. Zhang, "Live-SIMBA: an Imagej plug-in for the universal and accelerated single molecule-guided Bayesian localization super resolution microscopy (SIMBA) method," *Biomed. Opt. Express* **11**(10), 5842–5859 (2020).
32. S. Gao, "The source code of DETECTOR," Github, 2021), <https://github.com/GaoXiaoshan/DETECTOR>.